

# APPENDIX B

**Significant New Facts Revealed By Recent Meta Productions**

Prior Representation(s)	New Representation(s)
<b>LibGen Fiction 2024 Torrenting</b>	
<p><b>April 1, 2025:</b> “However, as made clear from a review of the source code files, the produced downloaded files list, and confirmed by my interview with the Meta engineer responsible for the torrent download, it is clear that none of the LibGen fiction library’s data was torrented by Meta in 2024.”<sup>1</sup></p> <p>“There is no evidence that the fiction portions of the LibGen library (separate from the torrent files) were downloaded during the dataset torrent download process that took place in 2024, and Ms. Wang confirmed to me that they were not.”<sup>2</sup></p>	<p><b>Sept. 12, 2025:</b> “We have retrieved the contents of the requested S3://fair-spark/xiaolan directory and have placed them on the source code computer for your review, and we are also producing a text file listing the files in that directory. We have recently learned that this directory contains files from an additional download of titles from the LibGen-Fiction dataset. Based on the dates of the files, this download appears to have taken place in early April 2024 and contains approximately 6,000 works downloaded from that dataset.”<sup>3</sup></p>
<b>LibGen Non-Fiction (“Sci-Tech”) 2024 Torrenting</b>	
<p><b>February 10, 2025:</b> “The first item in Table 2 shows that, for Libgen.rs Non-Fiction, only three copies of Plaintiffs’ works were downloaded. This does not represent three distinct books; it represents a single book (Ta-Nehisi Coates, <i>The Beautiful Struggle</i>), which was repeated three times in the downloaded dataset. No other Plaintiff works are contained in this dataset.”<sup>4</sup></p> <p>Both of Ms. Frederickson-Cross’s experts reports repeatedly state Meta torrented only a portion of LibGen Non-Fiction (Scitech) in 2024.</p>	<p><b>October 27, 2025:</b> “Meta has now conducted the follow-up referenced in our October 17, 2025 correspondence, and has identified additional responsive information. . . . We have now ascertained that the uptick in late May 2024 is attributable to a torrent download of files from the ‘Scitech’ portion of Libgen, executed by Ms. Wang from around May 29, 2024 until approximately June 2, 2024. . . . As a result of this newly discovered information, Meta anticipates producing additional documents as promptly as possible. Specifically, Meta intends to produce a list of the Libgen Scitech files that were downloaded as part of this late May 2024 download, as well as ‘parquet’ files that appear to have been generated from this download.”<sup>5</sup></p>

<sup>1</sup> Exhibit F, Frederiksen-Cross Second Rebuttal Report, ¶ 87 (Apr. 1, 2025).

<sup>2</sup> Exhibit F, Frederiksen-Cross Second Rebuttal Report, ¶ 91 (Apr. 1, 2025).

<sup>3</sup> Exhibit G, Letter from Phillip Morton to Plaintiffs’ Counsel (Sept. 12, 2025).

<sup>4</sup> Exhibit Q, Frederiksen-Cross First Rebuttal Report, ¶ 111 (Feb. 10, 2025).

<sup>5</sup> Exhibit S, Email from Meta Counsel to Plaintiffs’ Counsel, at 1-2 (Oct. 27, 2025).

Prior Representation(s)	New Representation(s)
<b>Guillame Lample 2022-23 Torrenting</b>	
<p><b>March 3, 2025:</b> “It was understood that Guillaume [Lample] had done work with LibGen, but we have not been able to track down any evidence of what had been downloaded or been able to find that -- any more details.”<sup>6</sup></p>	<p><b>Oct. 17, 2025:</b> “With respect to Mr. Lample, Meta produced tens of thousands of log files regarding the torrent files he downloaded.”<sup>8</sup></p>
<p><b>March 6, 2025:</b> “Another employee – I believe his name was Guillaume -- but I wasn't sure I had spelled that right or caught it right – might have used BitTorrent. He wasn't sure. Nikolay was not sure whether this individual had used BitTorrent, and the individual was no longer an employee, and there was no record of him having used it that I have seen; so that was a kind of a dead end.”<sup>7</sup></p>	<p><b>Sept. 26, 2025:</b> “Meta has possession of Meta-issued computers used by Mr. Lample. We reiterate that Mr. Lample is not a custodian and additional information from him is not relevant nor proportional to any current disputes.”<sup>9</sup></p>

---

<sup>6</sup> Exhibit A, Mike Clark 30(b)(6) (Torrenting) Tr. at 54:22-25 (Mar. 3, 2025).

<sup>7</sup> Exhibit R, Barbara Frederiksen-Cross Tr. at 216:3-11 (Mar. 6, 2025).

<sup>8</sup> Exhibit S, Email from Meta Counsel to Plaintiffs’ Counsel, at 13 (Oct. 17, 2025).

<sup>9</sup> Exhibit S, Email from Meta Counsel to Plaintiffs’ Counsel, at 47 (Sept. 26, 2025).

Prior Representation(s)	New Representation(s)
ML Hub/AIDC	
<p><b>Sept. 19, 2024:</b> “Further to Plaintiffs’ request, Meta identifies the following data sources within its possession, custody or control that have been searched and/or collected from to identify potentially responsive or relevant documents or ESI:</p> <ul style="list-style-type: none"> <li>• Email</li> <li>• Workplace Chat</li> <li>• Contracts</li> <li>• Google Suite</li> <li>• Workplace Groups</li> <li>• Workplace Posts</li> <li>• Workplace Notes</li> <li>• Cooley LLP email</li> <li>• Clearly Gottlieb email</li> <li>• Public Meta websites (e.g., research.facebook.com, ai.meta.com)</li> <li>• Research Reviews</li> <li>• Single Review Tool (SRT)</li> <li>• GitHub</li> <li>• Research Super Cluster (RSC)</li> <li>• S3 (AWS servers)”<sup>10</sup></li> </ul> <p><b>Oct. 21, 2024:</b> “Meta searched for and/or collected responsive documents in the data sources within Meta’s possession, custody or control that would have potentially responsive or relevant documents, including both non-custodial files and custodial files . . . Meta is not aware of any types of data sources with potentially responsive or relevant documents that were not searched.”<sup>11</sup></p> <p><b>Dec. 9, 2024:</b> “To be clear, Meta has collected and produced from custodial and non-custodial sources when identified during Meta’s reasonable search, including cases where no ESI custodian was involved.”<sup>12</sup></p>	<p><b>Oct. 17, 2025:</b> “Plaintiffs have known about AIDC prior to the December 13, 2024 close of discovery. <i>See e.g.</i>, November 13, 2024 Deposition of Michael Clark at 195:5-6 (“At that point, the dataset review process went from the tracker into a tool called AIDC.”). Thus, any disputes regarding collection of materials from AIDC should have been raised by December 20, 2024; Plaintiffs failed to raise this issue and it is thus waived. Additionally, Dkt. 470/616 required Meta to run a search term on emails and Workchats of specific custodians; it did not require Meta to run those searches on non-custodial sources. Notwithstanding the waiver and the request clearly falling outside of Dkt. 470/616, upon request from the Plaintiffs and as an attempt to resolve outstanding issues, Defendants agreed to and did produce the AIDC entries for LibGen and Anna’s Archive.”<sup>13</sup></p>

<sup>10</sup> Exhibit M, Email from Meta Counsel to Plaintiffs’ Counsel (Sept. 19, 2024).

<sup>11</sup> Exhibit E, Email from Meta Counsel to Plaintiffs’ Counsel, at 24 (Oct. 21, 2024).

<sup>12</sup> Dkt. 321, Meta’s Position, Joint Letter Brief re Plaintiffs’ Additional Discovery Requests (Dec. 9, 2024).

<sup>13</sup> Exhibit S, Email from Meta Counsel to Plaintiffs’ Counsel, at 11 (Oct. 17, 2025).